DISCLAIMER NOTICE

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

35177 A.O. 153 T-III-C

ZR-AP-061-16
MAY 1961

Mr. Koether has seen

262649

AD No.
ASTIA FILE COPY

SPACE-TIME SAMPLING AND FILTERING

Semi-annual Technical Documentary Note

Prepared For
AIR FORCE SPACE SYSTEM DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
Inglewood, California

ADVANCED RESEARCH PROJECTS AGENCY ORDER NUMBER 153-60
UNITED STATES AIR FORCE CONTRACT NUMBER AF 04(647)-644

GD
CONVAIR SAN DIEGO
Convair Division of General Dynamics Corporation

ZR-AP-061-16
MAY 1961

SPACE-TIME SAMPLING AND FILTERING

Semi-annual Technical Documentary Note

Prepared For
AIR FORCE SPACE SYSTEM DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
Inglewood, California

ADVANCED RESEARCH PROJECTS AGENCY ORDER NUMBER 153-60
UNITED STATES AIR FORCE CONTRACT NUMBER AF 04(647)-644

CONVAIR SAN DIEGO
Convair Division of General Dynamics Corporation

FOREWORD

The work documented in this report was performed by Dr. W. D. Montgomery and Dr. P. W. Broome of the Electronics Research Laboratory of Convair/Astronautics. It has been included as a part of the BAMBI study for reasons of administrative expediency in order to eliminate the necessity of the monitoring of two contracts although it was negotiated independently with different motivations and objectives.

This work was intended as a generalized study of the space-time filtering concept which is applicable to electro-optical detection systems. Thus, although it may be applicable to the BAMBI study, the technical approach and the results were not constrained by it and should not be interpreted within its context.

ABSTRACT

~~The present~~ This study is oriented toward enhancing the detection of a missile launch as seen from a satellite. Some preliminary considerations of satellite altitude, wave lengths viewed, and optical system used, show up the more important characteristics of the problem on the image plane of the optical system. Considering the intensity distribution on this plane as the input, three important spatial filters are developed, having two-dimensional outputs.

The first filter is a linear one whose simulation ~~has been carried out~~ on an IBM 7090, ~~The simulation uses~~ used a localized target against a random and non random background. The inputs and outputs are shown, ~~as taken from a Characteron tube.~~ The second filter, ~~is~~ a general statistical one, ~~which~~ has considerable discrimination ability but ~~which~~ needs a large amount of statistical information of the background. The third filter, ~~is~~ developed along the lines of decision theory, ~~and~~ turns out to be remarkably similar to the ~~previously mentioned~~ general one.

For each filter the application to multiple wave lengths as well as to point source and non point source targets is considered. The use of several filters combined in a serial or parallel fashion is mentioned.

CONTENTS

LIST OF FIGURES

1. INTRODUCTION

The immediate problem that the present study is concerned with is the detection of a missile launch as seen from a satellite. The satellite detection system will utilize data presented to it on an image plane and will absorb the data by way of an array of photo cells, a scanning photo cell or cells, a television-type scan, or some such similar technique.

The main emphasis of the study is on possible filtering actions that can be performed on the image plane intensity as an input. Such filtering will henceforth be carried out by what is termed a "spatial filter" whose design is based on certain a priori knowledge of the target and background. The purpose of the filter is to enhance the detection of the target. Filters which are discussed are of two types, the first type does processing in which the output is obtained by performing mathematical operations on the input. The second is a decision filter which votes according to a given criterion as to the absence or presence of a target in a given neighborhood of the image plane, thereby partitioning the output into two levels, absence or presence. Both types involve a mapping of the input plane onto the output plane.

The data which is seen at the image plane has already had some processing which must be considered in the analysis. This is dependent upon the wave lengths viewed, the optical imaging system employed, and the satellite altitudes considered. These are all external variables which determine some of the general properties of the incoherent light intensity distribution on the image plane of the optical system. The effect of these variables is to place constraints on the intensity distribution in the image plane.

At launch the dimensions of the plume of an Atlas missile about match those of the missile itself (80 feet by 10 feet). Thus from a satellite the missile at launch would most likely appear as a point source in the infrared. Out of the atmosphere the plume can be expected to increase its dimensions by at least ten times, and this combined with its possible decreased distance makes it no longer a point source. This addition of structure to the target as seen in the optical image is of importance as more complex filter criteria may be used to advantage. In general, structure is always present when the time dimension is included, even for a spatial point source if its velocity is not zero.

This point may be illustrated as follows: In Figure 1 the usual Raleigh criterion for resolution gives the resolvable dimension, r, as a function of the distance, D, for the two wave lengths--λ_i = 3.5 microns (infrared) and λ_v = .5 microns (visible). The objective aperture is assumed to be about one meter.

Since the plume radiation is most intense in the infrared (2.78μ and 4.2μ), this and a typical visible wave length have been chosen as being representative.

The resolution distance available on the image plane is determined by the wave length used as well as the aperture and focal length of the optical system. For an aperture of one meter and a focal length of four meters, the two wave lengths give resolution distances of .02 mm for 3.5μ and .002 mm for .5μ. Even though a scanning (say 100 lines per mm) of the image plane may not be able to take full advantage of this resolution the two cases previously mentioned for the target will still persist, namely the point source and the finite size.

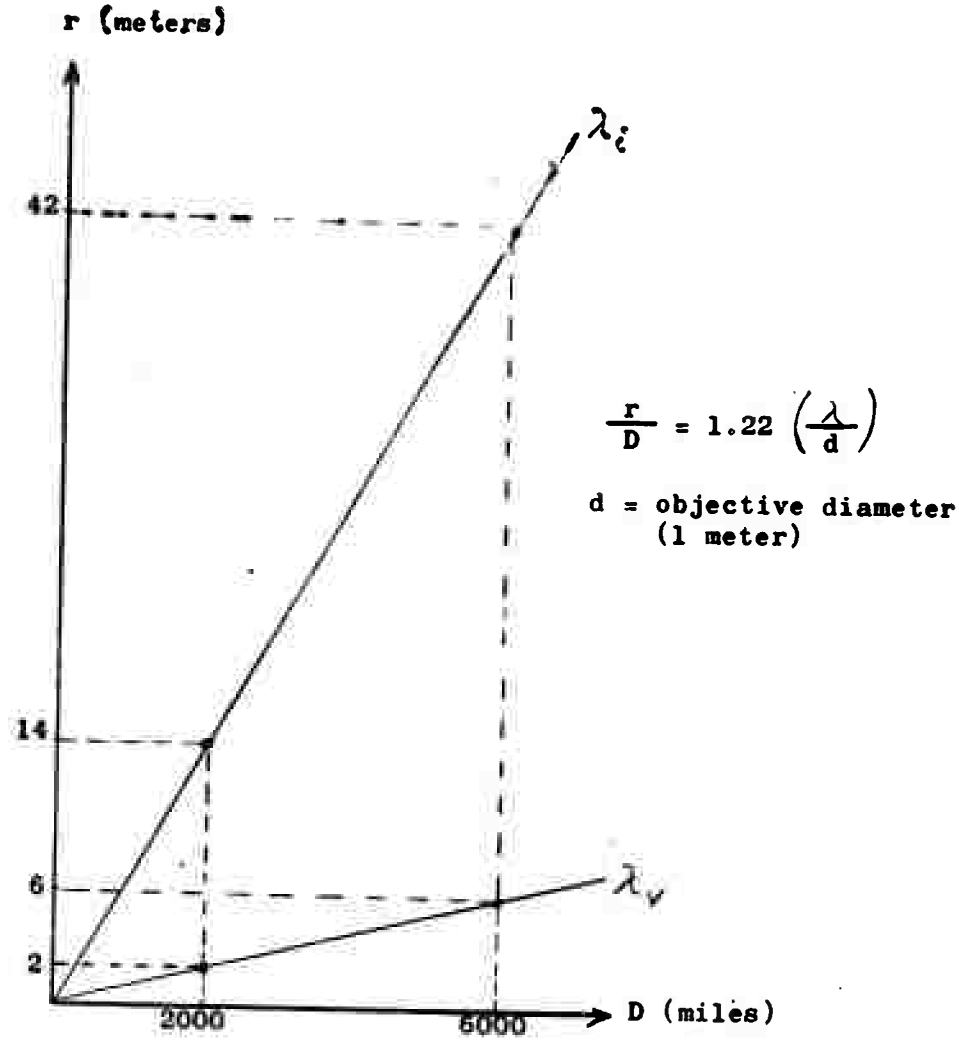


Figure 1
Resolution Distance vs Altitude

1.1 INFORMATION CONSIDERATIONS

The input to the spatial filters is a two-dimensional intensity distribution on the image plane of an optical system. Some remarks are in order concerning the output of such a filter. In particular, if the filter is considered as an information channel it would be desirable to keep the information loss[1] (equivocation) down unless considerable data reduction is anticipated within the filter itself. If the dimension of the output were lower, say one-dimensional with intensity a function of time, then it can be shown that there would be either considerable information loss or loss of continuity between input and output. The best that such a filter could do to minimize information loss alone would be to establish a one-to-one correspondence between the input plane and the output line, which can be done.

A topological theorem tells us that in such a case the condition of bicontinuity cannot hold simultaneously. This implies for the present circumstance that a localized target on the plane could not produce a localized output on the line and this would be a serious loss for detection systems. This one-to-one correspondence would involve no information loss

in a technical sense.[1] The only way that some degree of continuity could be maintained would be to make the mapping a many-to-one type (e.g., see Figure 2). This means that a localized target on the plane would produce a more localized output on the line but many input points would be mapped onto the same output point. This would imply that many distinct inputs would produce the same output giving rise to an information loss.[1] From the above considerations it would appear desirable to have a two-dimensional output for the spatial filters. Most of the filters considered will be of this type.

This is not meant to imply that the data must be processed in a parallel fashion. Serial processing with retention of dimension does not constitute an information loss process.

1.2 CONTINUOUS AND DISCRETE DESCRIPTIONS

The intensity distribution on the image plane is of a continuous nature but due to the resolution limitation it is actually discrete and in fact finite. One way this can been seen is by a consideration of the sinusoidal intensity waves that add up to form a two-dimensional picture.[2] This is its Fourier description and usually involves the summation of

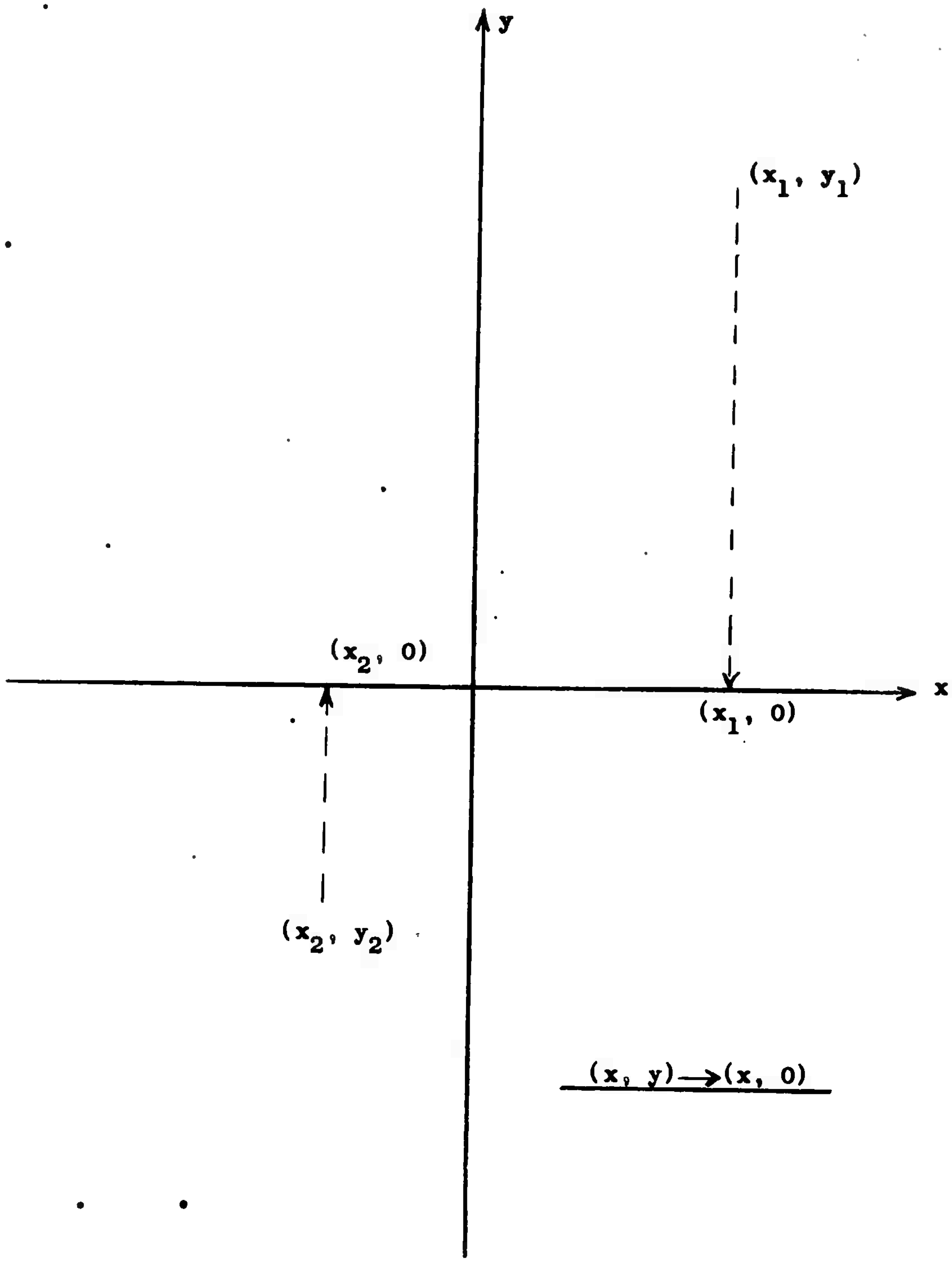


Figure 2
Localization in a Many-one Map

an infinite number of the sinusoidal waves. Since in the present situation there is a finite resolution distance, Δ, on the image plane there is no need for waves in the Fourier synthesis of wave length $< 2\Delta$. Such waves can only aid in the description of detail not found on the image plane. Thus the Fourier series is finite and we say the picture is band limited with an upper frequency limit (cycles/cm.) of $W = \frac{1}{2\Delta}$. Shannon's sampling theorem[3] applied to two dimensions shows how the picture can be completely described from a knowledge of the intensities at the points of a finite square grid of separation spacing Δ. The information at the intersections on this grid may replace the input plane of a spatial filter. This result correlates well with practical considerations since a scanning mechanism will no doubt be used to record the input data and such mechanisms are obviously limited to so many lines per mm. Similar arguments show that the output plane may be replaced by a grid of mesh no finer than that of the input grid.

1.3 THE NEIGHBORHOOD MODIFICATION PROCESS

Assuming that the input and output grids are superimposable, the general form of the spatial filter can be derived from a

few simple considerations. First of all since we are interested in the detection of a localized object, positional information should not be lost by the filtering action. This is assuming, of course, that the size of the target is relatively small in comparison to the whole field of view. Secondly, any statistical description of the background will show that correlation holds for relatively small distance, again in comparison to the whole field. By this we mean that the intensity level at one place on the image plane is not highly correlated with that at another place if they are very far apart. These statements will be made more precise in the section on filters where mathematical notation is introduced.

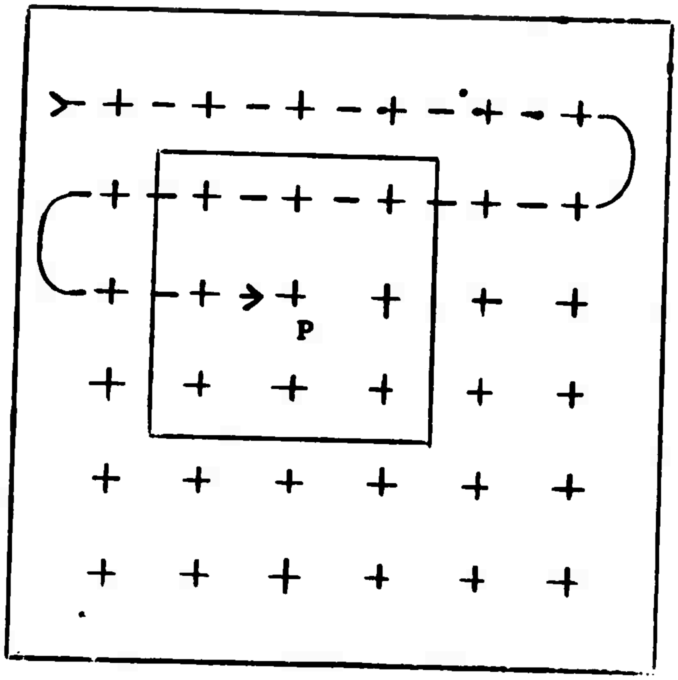
From these two observations it seems clear that the connection between the two grids should be as follows: The intensities in the neighborhood of an input grid point, P, should alone determine the intensity at the output grid point, P', where P' corresponds to P under the assumed superimposability. This mapping of a neighborhood of input intensities onto an output intensity is termed the "neighborhood modification process" (N.M.P.). The size of the neighborhood should be large enough

to contain the target as well as distances over which there is significant correlation in the background.

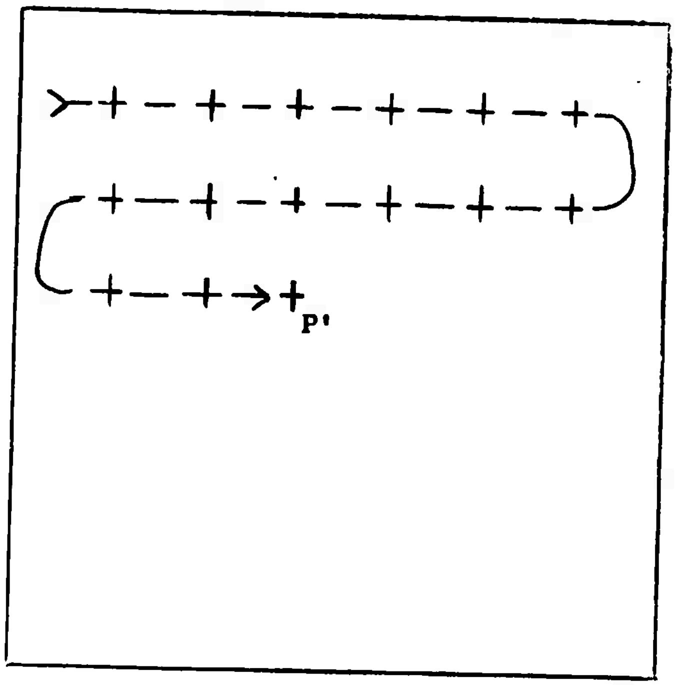
Thus the total filtering action can be described by a function $f(\underline{x})$ where $\underline{x}$ is an N-tuple vector whose components are the intensities at the points of the grid ordered in some way in the neighborhood of the point, P. For convenience the neighborhood will be chosen in the form of a square with P at the center and the points will be ordered from left to right and top to bottom. Thus, $f(\underline{x}) = y$ is the intensity at P' on the output grid. This is illustrated in Figure 3.

1.4 MULTIPLE WAVE LENGTH PROCESSES

In the combustion process of a rocket plume, as in most steady state burnings, there is a definite correlation between the infrared emission at one place in the plume and the visible emission at another place. This cross-correlation of intensities from different wave lengths is in addition to correlations occurring within any one wave length. This cross-correlation will also be present in the background (earth and cloud radiation) and thus additional discrimination can be gained by the use of multiple wave length spatial filters. The incorporation



Input Plane
(neighborhood outlined)



Output Plane

Figure 3
Neighborhood Modification Process
The intensities in the neighborhood are transformed and mapped to a single point in the output plane.

of more than one wave length into the scheme does not increase the theoretical difficulties by very much but in the implementation it would undoubtedly pose additional problems. For instance, one optical system might alternately illuminate two different input planes, each coated with a photosensitive material responding to that particular wave length.

The infrared and visible radiations tend to complement each other as to resolution and intensity. Most of the plume radiation is in the I.R. but the resolution is considerably better in the visible.

1.5 RETICLE SYSTEMS

There is a collection of mechanical devices which are used to detect the presence of a point source against a background of clouds, etc. These reticle systems, as they are called, employ a "chopping" action which usually consists of a collection of straight edge stops moving normal or parallel to their edges. This motion produces an off and on switching over the total field according to some pattern. Figure 4 illustrates a simplified rotating reticle. Some analysis of these devices has been done in the Fourier and Laplace planes.[4] It is clear from an examination of the image plane itself that their

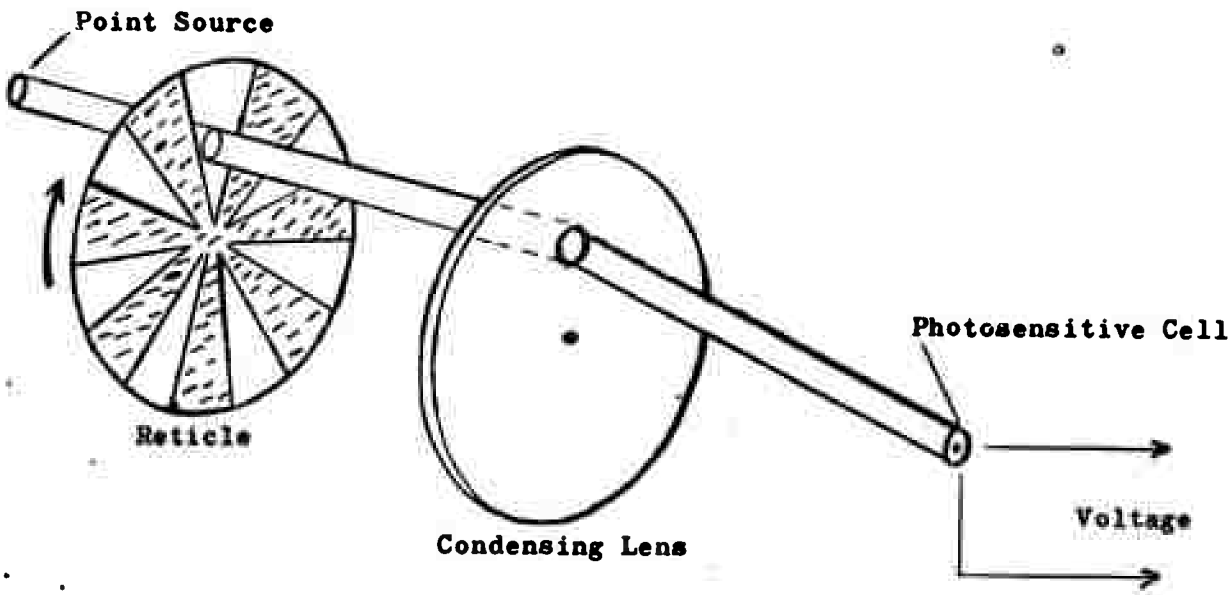


Figure 4
Rotating Reticle

utility lies, in the ability to pulse on and off a point source intensity while at the same time tending to suppress such pulsing of high intensity gradients that appear along edges whose average radius of curvature is not too small. The detection of the point source is thus reduced to that of detecting a pulsation in the integrated output intensity from the reticle. A clear advantage of most of these systems is their mechanical simplicity. Two serious defects are their inability to take advantage of much a priori knowledge concerning the background the their inability to supply good positional information unless high scan rates are used. The generalization of reticle systems to include non point sources is of questionable value and would tend to reduce to ordinary template matching. The spatial filters discussed in this report can utilize considerable a priori background information and detect non point sources as well as point sources. They, of course, will also tend toward more complexity.

2. DISCUSSION

There are a variety of ways to proceed toward the design of a spatial filter and all of these depend upon a priori choices of the mapping function $f(\underline{x})$, the optimization criterion, and the background and target descriptions. For each filter which is discussed, the single as well as multiple wave length cases will be considered. Also the targets will be both point source and non point source types with the latter being of fixed shape or having a statistical distribution of shapes. The optimization criterion employed is that of maximum transmitted peak target to r.m.s. background. One reason at least for using this criterion is that in the linear filter it makes use of signal shape whereas other criteria such as Weiner's do not.[5]

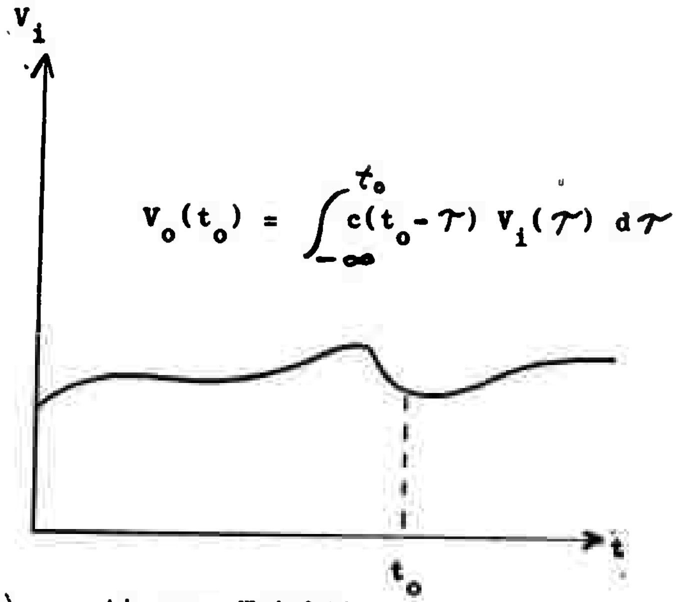
The free variables that are left after the above constraints are imposed are the form of the mapping function $f(\underline{x})$ and the background description. These have a complementary relationship such that if the mapping function is completely prescribed, as in a polynomial filter, the background description necessary is fixed. If the background is completely described statistically, as in the general filter, then the form of $f(\underline{x})$ is determined. These two situations as well as a decision filter will be described in what follows.

2.1 POLYNOMIAL FILTERS

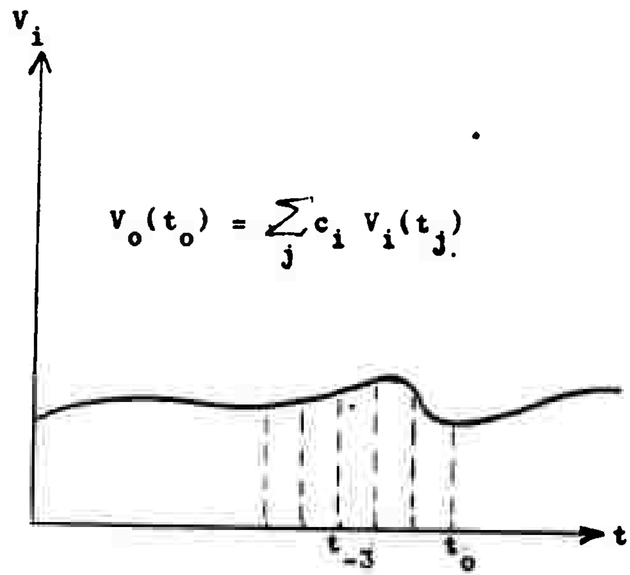
2.1.1 The Linear Filter

In an electronic linear filter whose input and output are voltages, the voltage at time t_o is determined by a linear weighting of the input prior to time t_o. This continuous weighting is performed by an integral operation. If a finite discrete weighting is desired this can be performed by a summation of past values using a fixed set of weighting coefficients. These two cases are shown in Figure 5. This operation corresponds to either a delay line filter or to a sampled data digital filter. In the case of spatial filters the discrete weighting procedure has complete two dimensional freedom in contrast to the restricted one dimensional freedom of the previous filter.

Figure 6 illustrates the necessary and sufficient form of $f(\underline{x})$ in order that the spatial filter be linear for the case of a 5 x 5 neighborhood. A linear filter is one in which a linear combination of inputs produces as an output the same linear combination of the corresponding individual outputs. The signal-to-noise ratio for the linear filter is:



a) tinuous Weighting



b) Discrete Weighting

Figure 5
Electronic Linear Filter

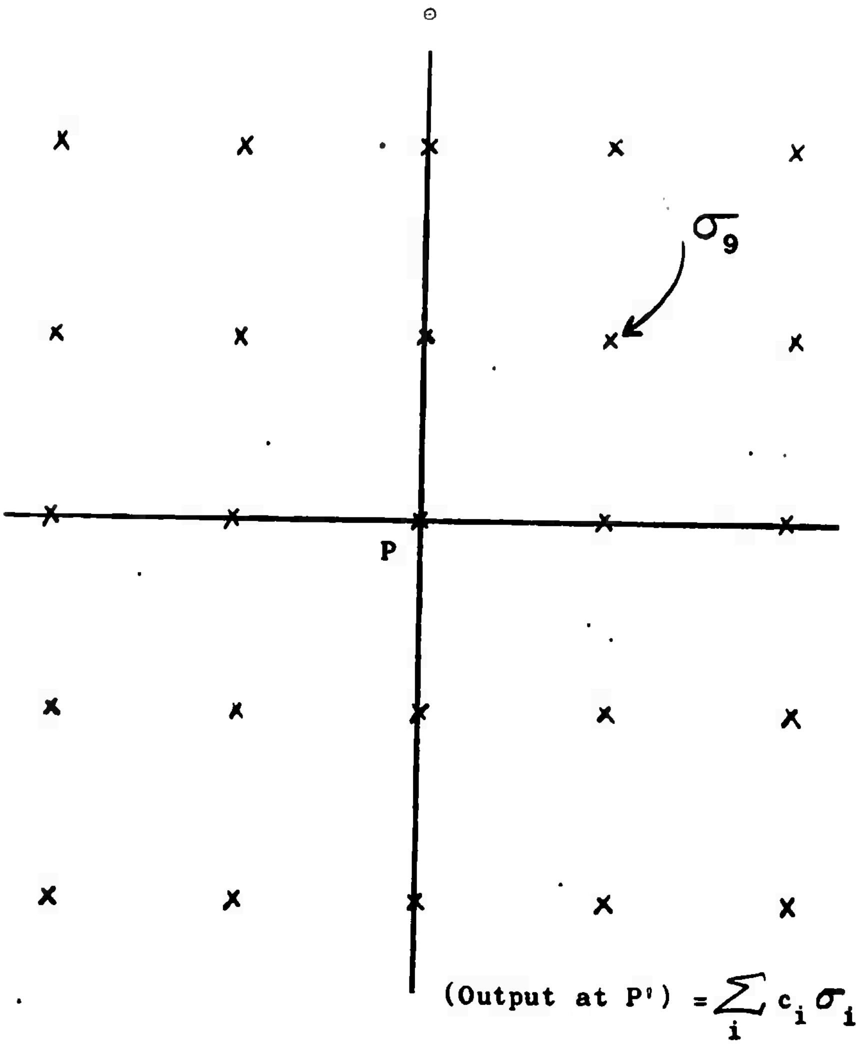


Figure 6
Spatial Linear Filter

$$\frac{\left|\overline{\sum_i c_i(s_i+n_i)}\right|}{\sqrt{\overline{\left(\sum_i c_i n_i\right)^2}}} \tag{1}$$

It is simpler to maximize the square or

$$R = \frac{\left[\sum c_i(s_i+\bar{n}_i)\right]^2}{\overline{\left(\sum_i c_i n_i\right)^2}} \tag{2}$$

where the summations are carried out over the neighborhood square and the bar denotes an average carried out over an ensemble of inputs. The c_i are the unknown weighting coefficients, the s_i are signal (target) values, and the n_i are noise (background) values. Here the signal is assumed to have a known shape and in fact peaks at the point P, the center of the neighborhood. The denominator of R can be written as $\sum_{ij} c_i c_j \varphi_{ij}$ where $\varphi_{ij} = \overline{n_i n_j}$ is the autocorrelation of the background. $\Phi = (\varphi_{ij})$ is a positive semidefinite matrix. So far two averages are indicated, $\overline{n_i} = \varphi$ and $\overline{n_i n_j} = \varphi_{ij}$ so that some comment concerning the ensemble mentioned above is in order. All of the spatial filters treated in this report are potentially adaptive. This means that the parameters

involved in the mapping function $f(\underline{x})$ (in the linear case the c_i) can be changed from time to time as new information about the background is acquired. This involves an assumption of stationarity in the background statistics over a period of time which includes that time used in recording an ensemble of background inputs and taking appropriate averages as well as the time during which the filter uses this information to enhance its detection ability. Such an assumption is always inherent in any statement concerning the use of a priori background data. If a sharp discontinuity in the background statistics occurs at some time then it is clear that stationarity does not hold and any adaptive system would tend to have a "blind spot" for a short time unless corrected by some additional memory scheme. Such a thing would be partly true for any human observer and so the difficulty cannot be completely overcome.

For the present linear case there are devices that will compute the autocorrelation of a two-dimensional input and which involve no moving parts. Eppler and Darius[6] have devised such an autocorrelator which is shown in Figure 7.

The A and B are two transparencies of the same picture where A is uniformly illuminated from the left so that the intensity

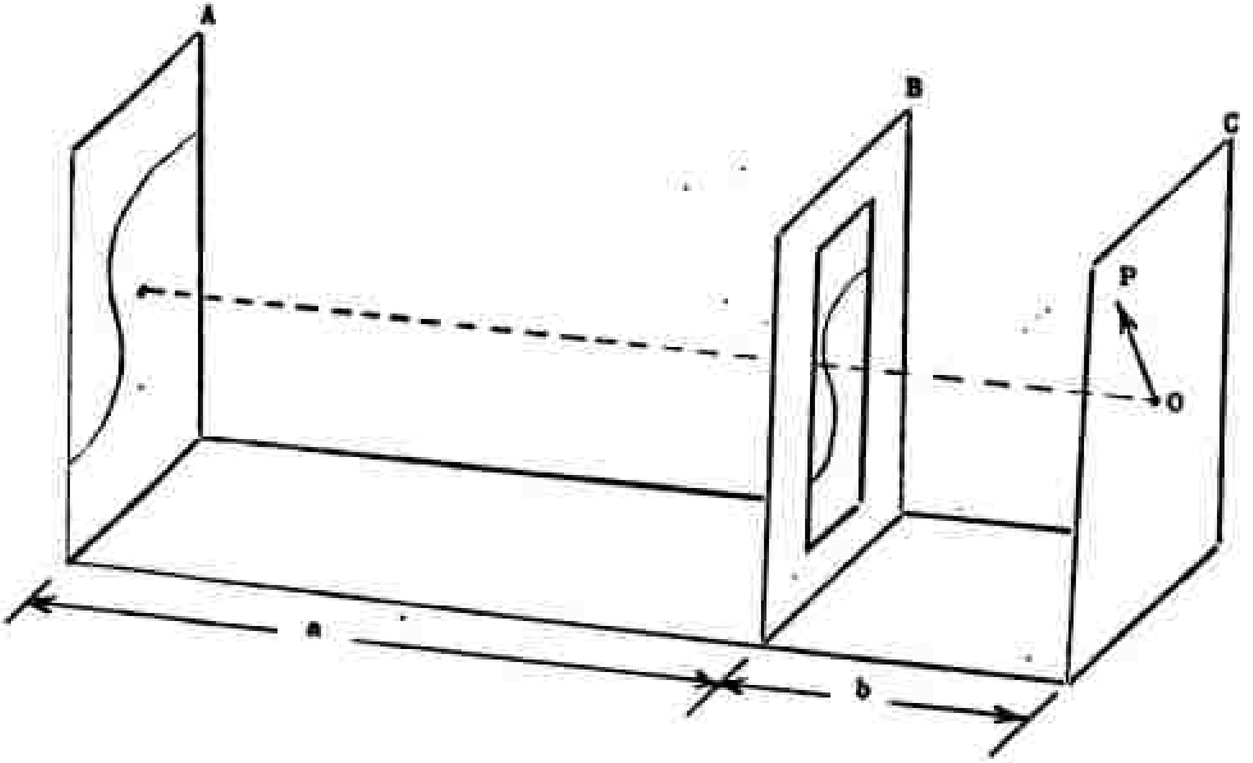


Figure 7
Autocorrelator

reaching point P on screen C by way of B is proportional to the autocorrelation for the vector <u>OP</u>. This depends upon two things:

(a) The spacings "a" and "b" must be such that the scaling ratio between transparencies is $\frac{a + b}{b}$.

(b) The geometrical fact that all straight lines intersecting the set of three parallel planes A, B, C are cut proportionately.

One shortcoming of this one as well as others[7] is that they need at least one transparency of the input for their operation. If this obstacle can be overcome the optical autocorrelator would appear to be a fairly simple device. Digital computer methods for computing autocorrelation have been found to be easily instrumented and are available as an alternate method.

Returning to equation (2) in an attempt to maximize the ratio R, there are two exceptional cases that should be mentioned first. There might conceivably exist a non trivial set of c_i such that $\sum_{ij} c_i c_j \varphi_{ij} = 0$, which would imply of course that Φ is not a positive definite matrix. In such a case the "best" choice of the c_i would be one of this type. Another

exceptional case is for R to be unbounded for non trivial c_i; here the choice on the c_i is unlimited. It is assumed in what follows that the exceptional cases do not occur so that $\underline{\Phi}$ is a positive definite matrix and an absolute maximum for R must be found among the solutions to the equations.

$$\frac{\partial R}{\partial c_n} = 0 \qquad n = 1, 2, \ldots \tag{3}$$

This procedure is carried out in Appendix A with the resulting vector solution.

$$\underline{\Phi}\, \underline{k} = a\, \underline{\sigma} \tag{4}$$

The $\underline{k}$ is the coefficient column vector consisting of the critical c_i, the $\underline{\sigma}$ has components $\sigma_i = s_i + \varphi$, and "a" is an arbitrary constant. Since $\underline{\Phi}$ is assumed positive definite it is non singular and

$$\underline{k} = a\, \underline{\Phi}^{-1}\, \underline{\sigma} \tag{5}$$

It is shown in Appendix A that a necessary and sufficient condition that this represent an absolute maximum is that the matrix $\alpha \underline{\Phi} - \underline{\sigma}\, \underline{\sigma}'$ be positive semi-definite. Here $\alpha = \underline{\sigma}'\, \underline{\Phi}\, \underline{\sigma}$ and primes denote the transpose.

Solution (5) represents only one value of R since R is a homogeneous function of the c_i of degree zero.

In the case of a point source all of the s_i are zero but the central value corresponding to the center of the neighborhood square. Thus the size of the square in this case is determined solely by background correlation. Independent of the fact that a complete specification of the form of $f(\underline{x}) = \underline{c}'\underline{x}$ demands a knowledge of the autocorrelation $\underline{\Phi}$ of the background, we can speak of a statistical measure of background correlation. Such a measure is necessary if one is to choose the size of the neighborhood square in an intelligent way.

Let $p_1(a:\underline{x},b)$ denote for the background the probability that given the free vector $\underline{x}$ with the intensity at its initial point in the interval between b and b + db, that the intensity at the terminal point is in the interval between a and a + da. If $p_2(c)$ denotes the probability that the intensity at any point is in the interval $(c, c + dc)$, then it is clear that $p_1(a:\underline{x},b) \approx p_2(a)$ for all a and b indicates no correlation for the vector separation $\underline{x}$. For each direction there would be a smallest length vector satisfying the above relation and the collection of these vectors will prescribe the smallest neighborhood square necessary to take full advantage of the correlation. In the present linear filter case as well as in other

non linear cases a simpler measure of the loss of correlation is for $\overline{n_i n_j} = \varphi_{ij} \approx \varphi^2 = (\overline{n_i})^2$ for sufficiently separated points i and j.

To incorporate multiple wave lengths into the present scheme would no doubt require a considerable addition in instrumentation but for the theoretical formalism only a minor notational change is needed. Let us consider for simplicity the case of two wave lengths λ_1 and λ_2. Of course one way to do this would be to design two separate linear filters and consider their output in a parallel fashion. The drawback here would be that no use is made of cross-correlation in the two wave lengths. So we will proceed as follows: Let the index, i, take on values from 1 to 2N where $N = m^2$ denotes the number of points in a neighborhood square of m by m size. The values i = 1 . . N will refer to wave length λ_1 and the values i = N + 1 . . . 2N will refer to wave length λ_2. Such an arrangement is both necessary and sufficient for the filter to be linear over both wave lengths. Thus for example $\overline{n_i n_j}$ for i = 1 and j = N + 2 gives the cross-correlation between the intensity in λ_1 at position one and the intensity in λ_2 at position two. It should be noted that the output of this filter is still in a single plane grid.

2.1.2 The Quadratic Filter

A polynomial filter somewhat more general than the previous one is a homogeneous quadratic one. In this case $f(\underline{x}) = \underline{x}'A\underline{x}$ is a quadratic form having a symmetric matrix A. If the same optimum detection criterion as was applied to the linear filter is applied to this one, the following ratio stands to be maximized:

$$R = \frac{\left(\sum_{ij} a_{ij}\, \sigma_i\, \sigma_j\right)^2}{\overline{\left(\sum_{ij} a_{ij} n_i n_j\right)^2}} \tag{6}$$

The subscript indices take on values from 1 to N, i.e., the numbering of points in a neighborhood square. The a_{ij} are elements of the matrix A while σ_i and n_i are the signal and noise intensities, respectively. The bar indicates an average over an ensemble of noise backgrounds. To maximize R the a_{ij} must satisfy the following equations:

$$\sum_{ij} a_{ij} \varphi_{ijkl} = b\, \sigma_k \sigma_l \quad \text{for } k, l = 1, \ldots N \tag{7}$$

The $\varphi_{ijkl} = \overline{n_i n_j n_k n_l}$ so that they represent a correlation in the noise at a quartet of points, i, j, k, l as distinct from the autocorrelation $\varphi_{ij} = \overline{n_i n_j}$ which represents a correlation in the noise at a pair of points, i, j. Since there are N^2

equations in the N^2 unknowns a_{ij} it is clear that in general there will be a unique set of a_{ij} determined and hence a unique function f. The "b" is an arbitrary constant which appears for the same reason as "a" in equation (4), namely because the ratios to be maximized are homogeneous of degree zero in the unknown constants.

The adaptation of the quadratic filter to a point source and multiple wave lengths proceeds along lines exactly parallel to those of the linear filter.

If one proceeds to non homogeneous polynomials or to ones of higher degree an increased knowledge of correlation in the noise is demanded. A more general and satisfying approach is to assume a somewhat complete knowledge of correlation in the noise and then proceed to find the function f which will maximize the signal-to-noise ratio.

2.2 A GENERAL STATISTICAL FILTER

For any ensemble of noise backgrounds consider the following function $\mathcal{V}(\underline{n})$, where the vector $\underline{n}$ denotes a noise intensity distribution over a neighborhood square (m by m). The image

plane has been replaced by an input grid so that the vector $\underline{n}$ comes from a Euclidean N-space ($N = m^2$). The function ν is defined as a probability density function for the random variable vector $\underline{n}$. This means that the integral

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_N}^{b_N} \nu(\underline{\xi})\, d\xi_1\, d\xi_2 \ldots d\xi_N \qquad (8)$$

gives the probability that an arbitrary neighborhood of an arbitrary input grid taken from the ensemble will have a distribution of intensities satisfying the following inequalities.

$$a_i < n_i < b_i \qquad \text{for } i = 1 \ldots N$$

where $\underline{n} = \begin{pmatrix} n_1 \\ \vdots \\ n_N \end{pmatrix}$

The above integral will be abbreviated by $\int_{\underline{a}}^{\underline{b}} \nu(\underline{\xi})\, d\underline{\xi}$ and it is clear that $\int \nu(\underline{\xi}) d\underline{\xi} = 1$ with infinite limits understood. The size of the neighborhood square is dictated by the area over which the noise has "significant" correlation properties. In practice where the ensemble is somewhat restricted the function ν can always be computed. In fact, regardless of the size of the ensemble the computed function represents the most complete statistical knowledge consistent

with that size. It should be noticed that from the function ν such functions as the autocorrelation and four point correlation descriptions of the noise can be derived.

Analogous to the function $\nu(\underline{n})$ another probability density function $\sigma(\underline{s})$ will be introduced. Assuming that the neighborhood square is large enough to contain the localized signal as well as noise correlation, $\sigma(\underline{s})$ will describe the probability with which the signal takes on various orientations as well as positions in the square. If several different signals are involved σ can incorporate these as well. More precisely the integral

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_N}^{b_N} \sigma(\underline{\eta})\, d\,\eta_1\, d\,\eta_2 \cdots d\,\eta_N \tag{9}$$

gives the probability that an arbitrary neighborhood containing a signal superimposed on an average noise background will have a distribution of intensities satisfying the following inequalities.

$$a_i < s_i < b_i \qquad \text{for } i = 1, \ldots N \quad ,$$

$$\text{where } \underline{s} = \begin{pmatrix} s_1 \\ \vdots \\ s_N \end{pmatrix} \quad . \tag{10}$$

With the above definitions of ν and σ it is clear that the average transmitted signal and the average square of the transmitted noise are respectively

$$\int f(\underline{\eta})\,\sigma(\underline{\eta})\,d\underline{\eta} \quad \text{and} \quad \int f^2(\underline{\xi})\,\nu(\underline{\xi})\,d\underline{\xi} \tag{11}$$

Thus the ratio to be maximized is

$$R = \frac{\left(\int f(\underline{\eta})\,\sigma(\underline{\eta})\,d\underline{\eta}\right)^2}{\int f^2(\underline{\xi})\,\nu(\underline{\xi})\,d\underline{\xi}} \tag{12}$$

and the variational calculus is employed in Appendix B to find the unknown function f. Assuming $\varphi(\underline{\xi})$ to be the function sought the desired result is

$$\varphi(\underline{\xi}) = A\,\frac{\sigma(\underline{\xi})}{\nu(\underline{\xi})} \tag{13}$$

with A an arbitrary constant. The ratio in Equation (13) is intuitively correct since any intensity distribution in a neighborhood square centered at point P which has a much higher probability of representing a signal than noise will give a strong output intensity at P'.

In thinking of $\sigma(\underline{\eta})$ as a scaler point function in an N dimensional space, the point source target case corresponds to a σ which is only sizable for points in the space which have a large central coordinate value for $\underline{\eta}$.

Thus the computation of σ for the point source case depends only upon the average background and the statistical distribution of amplitudes of the source.

The multiple wave length case can be handled in a way similar to that of the linear filter. For the case of two wave lengths the dimension of the vector $\underline{\xi}$ is doubled so that the first half of the coordinates refer to intensities in λ_1 and the second half to intensities in λ_2. The ν and σ functions then incorporate all of the cross-correlation of intensities in λ_1 and λ_2.

Any instrumentation or simulation on a computer of this kind of a statistical filter requires a large memory capacity. Ways of obtaining approximate and useful estimates of the ν and σ functions are at the present time being worked out. The filter is somewhat of an ideal to be approached and also shows how the mapping function $f(\underline{x})$ is completely determined, i.e., to within a multiplicative constant, when the background description is given.

2.3 A DECISION FILTER

Using the same ν and σ functions as previously defined, it will be interesting to show how an entirely different approach

to the detection problem can lead to a result strikingly similar to that obtained in Section 2.2. Middleton[8] in applying the work of Wald[9] to the detection of signals in noise introduces a decision function $\delta(d_i : \underline{\xi})$. Here $\delta(d_o : \underline{\xi})$ gives the probability that no signal is present in a neighborhood square given the intensity distribution $\underline{\xi}$. Similarly $\delta(d_1 : \underline{\xi})$ gives the probability that a signal is present. From the definition it follows that

$$\delta(d_o : \underline{\xi}) + \delta(d_1 : \underline{\xi}) = 1 \tag{14}$$

The optimum criterion to be used in designing the decision filter is called that of an ideal observer[8]. This criterion minimizes the total average error of which the filter is capable. There are two types of errors. One of them, called the false alarm, gives an output of signal present when no signal is present. The other, called the miss, gives the output of no signal present when a signal is present. The average probabilities of these two types will be denoted by βq and αp, respectively. Here p and q denote the a priori probabilities of the presence or absence of a signal with $p + q = 1$. Thus α gives the average conditional probability of a miss while β gives the average conditional probability

of a false alarm. The relations giving α and β in terms of the decision function δ and the density functions ν and σ are as follows:

$$\begin{aligned} \alpha &= \int \sigma(\eta)\,\delta(d_0:\eta)\,d\eta, \\ \beta &= \int \nu(\eta)\,\delta(d_1:\eta)\,d\eta \end{aligned} \tag{15}$$

The ideal observer then seeks to minimize the total average error $T = \alpha p + \beta q$.

$$T = p \int \sigma(\eta)\,\delta(d_0:\eta)\,d\eta + q \int \nu(\eta)\,\delta(d_1:\eta)\,d\eta \tag{16}$$

Using Equation (14) the total average error becomes

$$T = p + \int \left[q\nu(\eta) - p\sigma(\eta) \right] \delta(d_1:\eta)\,d\eta \tag{17}$$

The choice of δ which minimizes T is clearly the one which has $\delta = 0$ when the integrand is positive and has $\delta = 1$ when the integrand is negative. Thus

$$\delta(d_1:\eta) = \begin{cases} 1 \text{ if } q\nu(\eta) - p\sigma(\eta) \leq 0 \\ 0 \text{ if } q\nu(\eta) - p\sigma(\eta) > 0 \end{cases} \tag{18}$$

or in more familiar terms,

$$\delta(d_1:\eta) = \begin{cases} 1 \text{ if } \dfrac{p\sigma(\eta)}{q\nu(\eta)} \geq 1 \\[2ex] 0 \text{ if } \dfrac{p\sigma(\eta)}{q\nu(\eta)} < 1 \end{cases} \tag{19}$$

In the language of decision theory the space of inputs has been

partitioned into two parts such that a yes or no decision can be associated with each one. Comparing this result with that of Equation (13) it can be seen that for $A = \frac{p}{q}$ the filter of Section 2.2 becomes a decision filter where an intensity at P' greater than one is interpreted as a signal present in the neighborhood square about P. If the intensity is less than one no signal is assumed present.

The adaptation of the decision filter to a point source and to multiple wave lengths is exactly the same as for the general statistical filter. The similarity of the final form of $f(\underline{x})$ for the last two filters, considering their divergent approaches, is somewhat remarkable and tends to give considerable weight to that form of the function.

2.4 COMBINING FILTERS

Each filter which is designed is designed for the purpose of distorting the optical plane data in some certain desirable way. The linear filter does this by improving the signal-to-noise ratio using signal description and noise autocorrelation. The decision filter evaluates the ratio of the probabilities of the input having occurred in the signal to its having occurred in the noise. An infinitude of filters can be designed which require statistical descriptions lying somewhere between these extremes.

Decisions can be made as to the probability of a target's presence or absence in the input optical plane dependent on the filter output. The decision filter makes this decision in its operation. Various types of decision criteria may be used. For example, the output intensity of any filter may be partitioned at some level, all levels above the partitioning point indicating signal presence, all points below indicating absence.

Filters which have a two-level output, indicating presence or absence, are called property filters. Their effect is to divide the universe of possible patterns into two classes, the 0 class and the 1 class. Property filters as applied to the thesis problem would, for example, include measurements of the presence or absence of the target in several spectral ranges.

Property filters may be designed on any reasonable basis. Their evaluation may of necessity be empirical. The term as used here means simply that the filter measures some property of the image and votes accordingly.

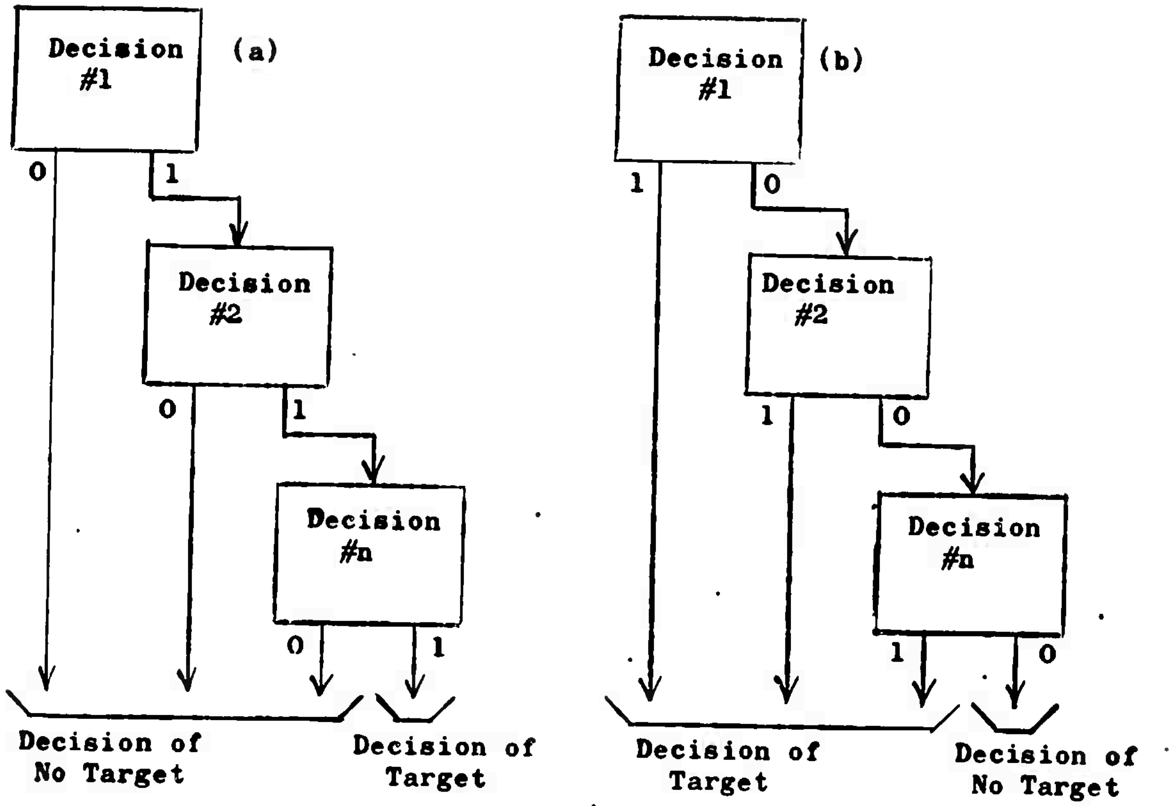
Several decision processes may be combined to enhance the accuracy of the overall decision process. It is probably easier to find a large set of properties, each of which gives a little useful information about the target, than to find a few which

are just right. This is especially so in the adaptive system.

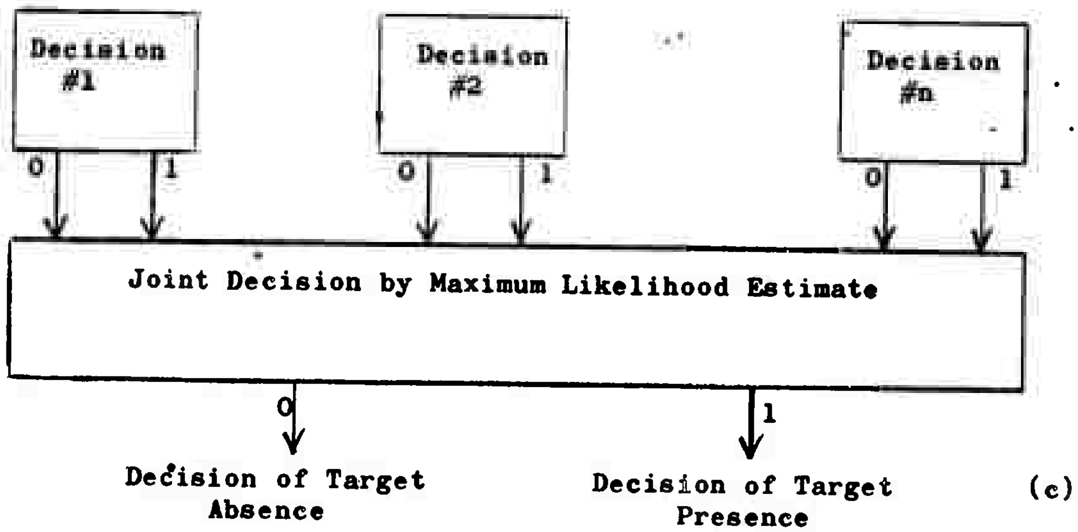
Two basic ways of combining decision processes are illustrated in Figure 8. The serial decision requires excellent discrimination ability at each step.

The serial scheme (a) minimizes false alarms at the expense of permitting misses. The serial scheme (b) minimizes misses at the expense of permitting false alarms. Either case would probably be intolerable in a system in which the machine has the final say. If the final decision were human, scheme (b) would be better as it minimizes misses. The parallel decision (c) does not require such tight requirements on the individual decisions but requires that they be combined by some good method in the final decision block. By good, it is meant that the outputs of highly discriminating decision blocks are weighted heavily while less reliable decision blocks are weighted lightly.

Each filter has associated with its decision ability a probability value of its decision being correct or incorrect. The purpose of a design which retains some residual adaptation is to improve the reliability of the decision as time goes on. Thus, the autocorrelation function in the linear filter is



Simple Serial Decision Schemes



Parallel Decision Scheme

Figure 8
Decision Filter Combinations

continually updated so that it remains representative of the noise. It seems clear that the property filters used in a detection scheme should be designed with the problem constraints in mind.

The problem of designing a system which uses these concepts breaks itself rather naturally into two parts. The first is to design property filters which are in some sense optimum and which, again in some sense, retain residual adaptability. The second is to use the decisions made by these property filters so as to improve the overall decision ability of the system. Each filter has four probabilities associated with it: P_{11} the probability that it decides on a target being present, given that a target is present. P_{10} the probability that it decides on a target being present, given that a target is not present.. P_{01} the probability that it decides against a target being present, given a target is present. P_{00} the probability that it decides against a target being present, given a target is not present. Ideally, P_{11} and P_{00} are each of value 1 and P_{01} and P_{10} are zero. The adherence to this optimum is the goal of the filter design. If this were possible for a single filter then there would be no need of using more than one.

The decision filter, sec 2.3, is a single filter which minimizes miss and false alarm decisions, P_{10} and P_{01}. Therefore it is ideal. The quantity of data which it must keep available, however, is prodigious and this is a limitation which may not be easily eliminated. For example, if it were to look at ten sequential images (time) of twenty x twenty points, and each point had a ten level number associated with it, the number of possibilities which can occur and which must be accounted for is 10^{4000}, i.e., 4,000 points taking on 10 levels independently. It is obvious that some compromises must be made in the instrumentation of such a filter which will tend to make it less than ideal.

The filters discussed in secs 2.1, 2.2, 2.3 of this report used in several spectral regions are optimum in the sense in which they were derived. They will nevertheless have success probabilities associated with them and if the design criteria are reasonable their success probabilities will be much larger than chance.

Likelihood estimates can be assembled from the P_{ij} probabilities that have been introduced. These are being considered at the present time as a method of combining decision filters in parallel.

2.5 SIMULATION

In order to develop some feeling for what can be attained by the use of different filtering and detection criteria, computer simulations of systems using these criteria are being established. This is a two purpose program. First, to see how well the mathematics applies, and second to obtain some idea of the requisite complexity of a system using these methods.

The hardware being used in the simulations is:

1. A picture scanner (still under construction) which partitions a photograph into 100 x 100 equally spaced points and generates a number for each point proportional to the average intensity over a circle centered on that point. The circles are situated so that they touch but do not overlap. These numbers are then converted into a ten-level digital code and placed on IBM magnetic tape.
2. An IBM 7090 high speed digital computer with a core storage of about 30,000 words. The picture contains 10,000 computer words so two pictures plus program fit easily into the core.

3. An output device, the SC 4020 microfilm recorder which prints a 100 x 100 point picture. The technique used is to partition the output field of an oscilloscope into 100 x 100 equal squares and to print a number of dots in each square dependent upon the grey level, i.e., a 7-level is 7 dots, a 3 is 3 dots, etc. The SC 4020 recorder utilizes a Characteron shaped beam tube. The output oscilloscope picture is photographed. Variations of this approach using the SC 4020 are under investigation to see what improvements can be made.

The manipulation of the picture data is dependent upon the computer programming. Commensurate with the computer capabilities, subroutines have been (and are being) written so that new processing ideas can be checked by writing a driver program which calls for subroutine as required. For example, the autocorrelation function is obtained for a picture by calling AUTOCOR and stating the core location of the picture. Printouts of a picture at any stage of processing is also available as a subroutine.

A numerical picture output is available when desired by using an IBM 720 high speed printer. A picture printout subroutine

has been written which gives a picture output of up to 100 x 100 points with a decimal number printed at each point to represent the level. For reasons of size compression, we have been normalizing and quantizing the output into grey levels of zero to nine.

2.5.1 Autocorrelation

The autocorrelation program evaluates

$$\phi(\underline{R}) = \frac{1}{M} \sum_{\underline{r}} F(\underline{r}) F(\underline{r}+\underline{R})$$

where $\underline{R}$ is the vector separation between points $F(\underline{r})$ and $F(\underline{r}+\underline{R})$, the intensities at the sampling points, and M is the number of such points considered. $\underline{r}$ is a dummy scanning vector. $\phi(\underline{R})$ is a measure of the statistical influence of the value at one point in a picture upon the value at a point separated by a vector distance $\underline{R}$.

Since $(\underline{r}+\underline{R})$ does not necessarily lie in the picture matrix, the computation prescribed is evaluated by considering only those values of $\underline{r}$ which permit $\underline{R}$ to terminate on points in the matrix.

The autocorrelation function at the origin, $\underline{R} = \underline{0}$, is the average square of the function. For arguments larger than the

distance over which one point has influence over another, the autocorrelation function approaches the square of the average value of the function.

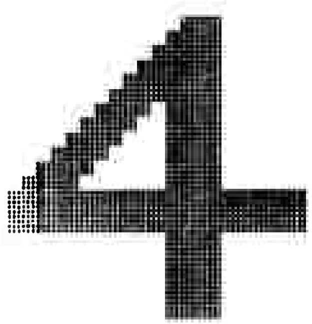
Figures 9 and 10 are examples of the autocorrelation function. The "a" picture is the input, $F(\underline{r})$. The "b" picture is the autocorrelation function, $\phi(\underline{R})$.

Figure 9, a number 4 and its autocorrelation function, demonstrates two properties of the autocorrelation function. One, the autocorrelation function is independent of translation of the image being transformed. Two, although structural information is lost, the distance over which correlation is important can be inferred from the image data. Note the strong correlation in the direction of the three lines making up the input image.

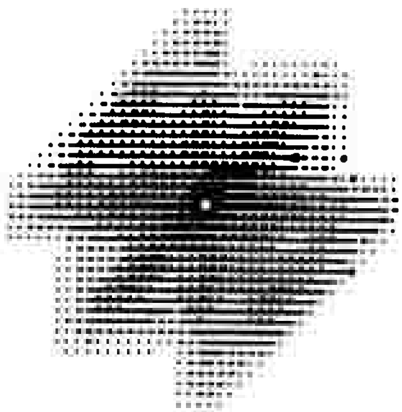
Figure 10, numbers between 0 and 9 taken from a random number table and its autocorrelation, illustrates the lack of statistical influence on one point over another when each point is completely independent.

2.5.2 Neighborhood Modification

The neighborhood modification linear weighting program performs a linear weighting of the picture intensity values on the input

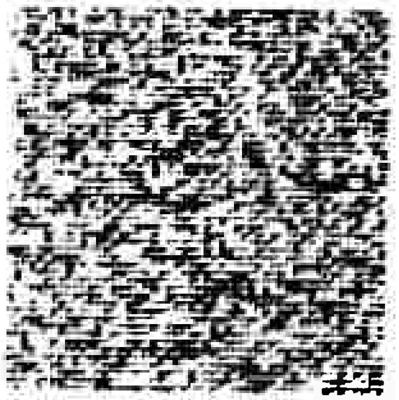


a) The numeral 4

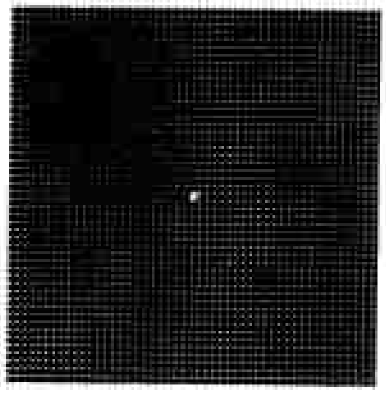


b) Autocorrelation of the numeral 4. The center value, shown as zero, is a ten level. The other levels are normalized to the center level.

Figure 9
Autocorrelation of Numeral 4



a) Ten level random noise.



b) Autocorrelation of (a). The center value, shown as zero is a ten level. The other levels are normalized to the center level.

Figure 10
Autocorrelation of Random Noise.

plane in a square of up to 25 x 25 points and plots the sum at a point in output plane at a point corresponding to the center of the square. It does this for every possible center on a picture of size up to 100 x 100 points. See Figure 3. Linear weighting of this type has a direct analogy with the one-dimensional linear filter. Since values on all sides of the point being processed are available, zero phase shift filters are possible.

2.5.3 Equivalence

The equivalence subroutine equates vector relations between points to matrix relations between entries.

For example, the autocorrelation between points with a separation of one horizontal element in the image plane, $\phi(1, 0)$, is the same for all i and j points of the neighborhood with this separation. There are $2n(n-1) + 1$ free and different autocorrelation vectors that will fit into an nxn neighborhood square.

2.5.4 Matrix Inversion

The matrix inversion subroutine solves the matrix equation,

$$\sum_i k_i \varphi_{in} = a\,\sigma_n$$

In the filter simulation, k_i are the filter constants to be found, φ_{in} is the autocorrelation of the noise between points i and n, and σ_n is the target description plus average noise. This is followed by a filter design subroutine which places the k's in the neighborhood modification scheme.

2.5.5 The Linear Filter Simulation

The optimum filter according to the criterion given in 2.1 was simulated as shown in Figure 11, with examples shown in Figures 12 and 13.

Pictures of size 28 x 28 elements were read into the computer by hand. 100 x 100 element pictures will be read in when the photo scanner is completed. The program solves for an optimum filter from noise and signal data to improve the signal-to-noise ratio for a signal + noise picture.

The output are normalized to the highest picture value and quantized into levels 0 to 9.

This filter is easy to instrument and an inexhaustive variety of modifications of the general idea are also possible without requiring much addition. In the detection filter maintenance of target shape was not included as a constraint. Thus, its only goal in life is to peak up at the point when it has a target centered in the neighborhood. It is seen to do this.

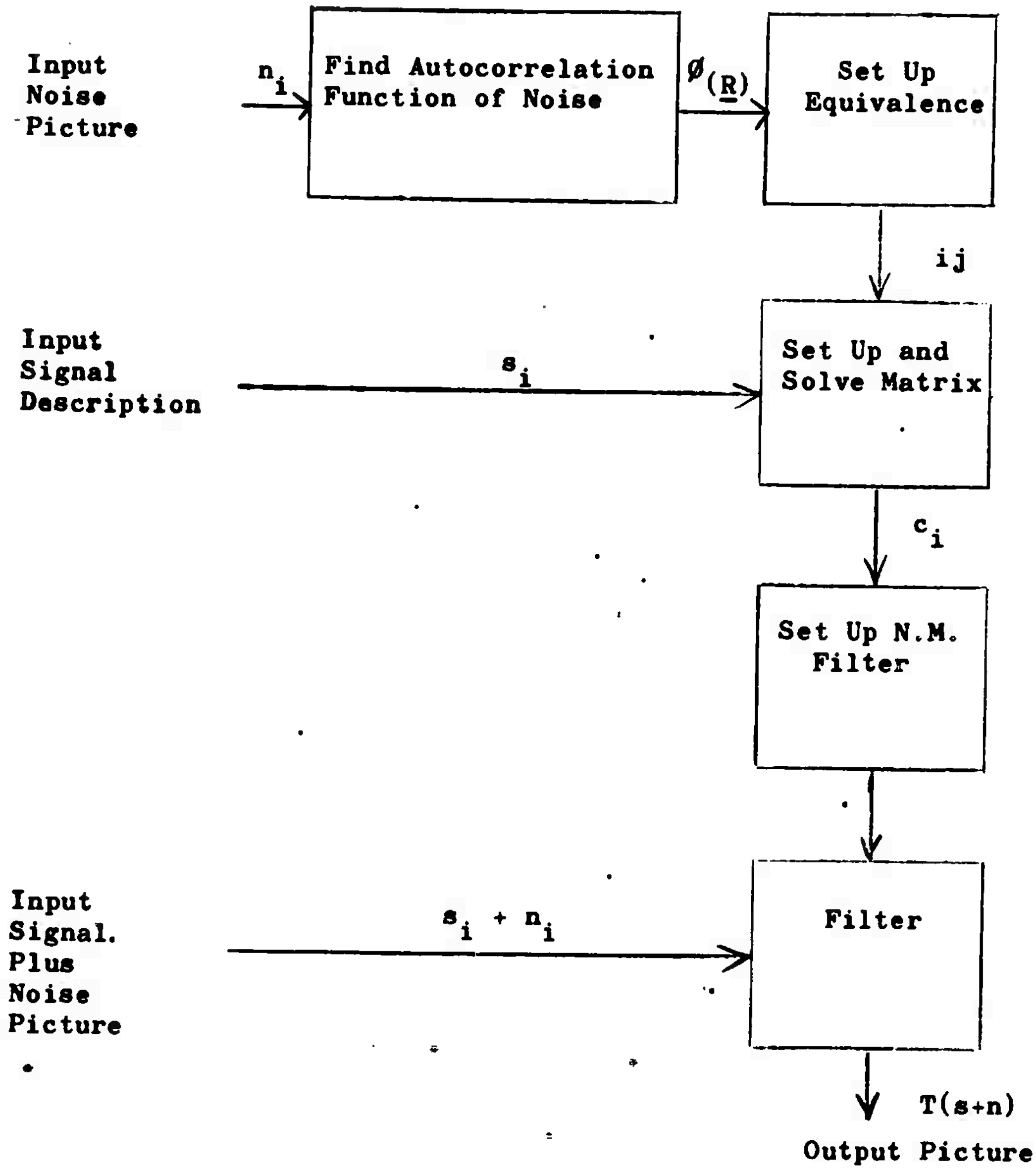
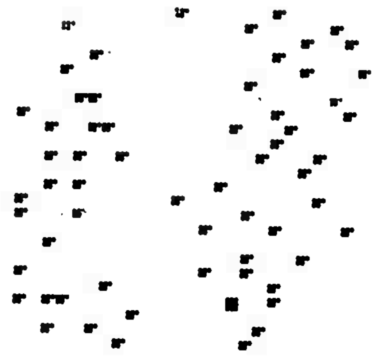
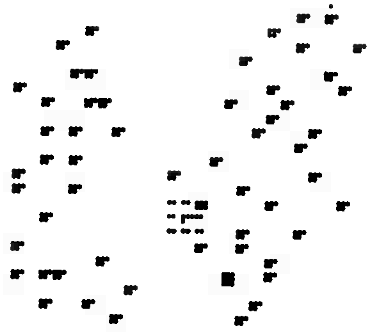


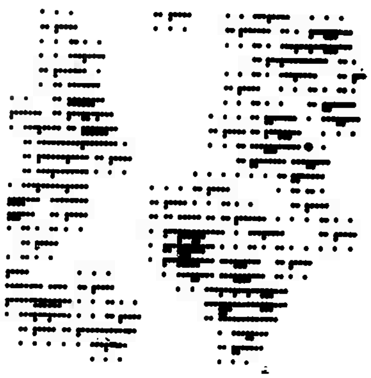
Figure 11
Computer Flow Chart for the Simulation of
the Linear Filter



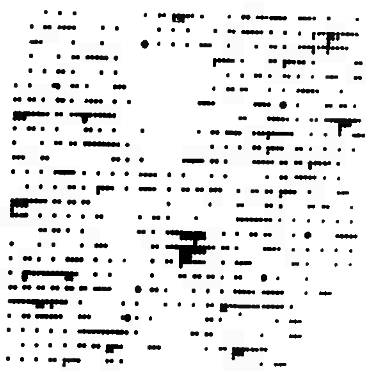
a) Dumbbell noise without target



b) Dumbbell noise plus target

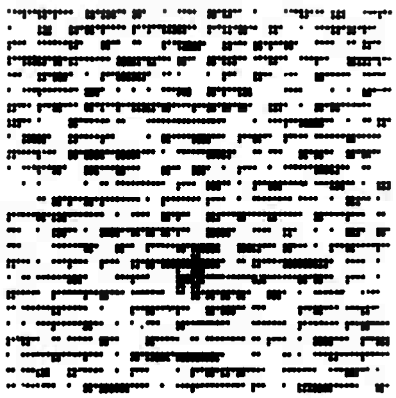


c) Dumbbell noise plus target filtered by a 3 x 3 neighborhood.

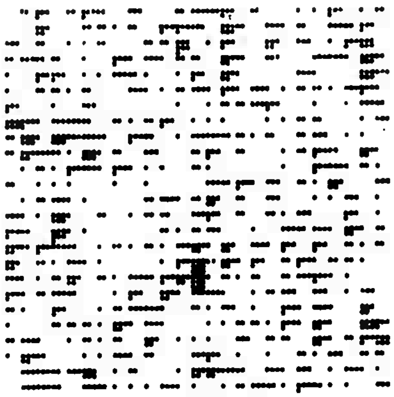


d) Dumbbell noise plus target filtered by a 5 x 5 neighborhood.

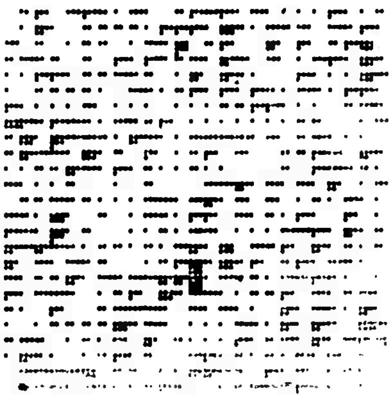
Figure 12
Simulated Linear Filter (Case 1)



a) Random numbers plus target



b) Random numbers plus target filtered by a 3 x 3 neighborhood



c) Random numbers plus target filtered by a 5 x 5 neighborhood

Figure 13
Simulated Linear Filter (Case 2)

The presence of a target on the image plane can be determined by either man or machine. This filter, which does not maintain structure of the target is probably better when using machine interrogation since human decision is generally based on structure. This will be investigated further when real data become available.

The target used was the same in the two cases which have been run. It consists of a 3 x 3 square with a center value of level 9 and peripheral values of level 4, i.e.,

444
494
444

It was placed in the same location in both cases, i.e., centered on the 13th point from the right and the 9th point from the bottom of the signal plus noise input picture.

For Case 1, Figure 12, the noise considered is dumbbell shaped. The dumbbells are oriented at an angle of 135° with respect to the positive x axis. Each dumbbell consists of two point source intensities of 10 level separated by a distance of $2\alpha\sqrt{2}$ (α is the spacing of the elements in the image plane). The dumbbells were placed in the image picture according to a random number table so that they occurred at a relative frequency of 1 in 20. Thus, on the average, one element in ten is sampling one of the sources.

In the computer simulation we let $\alpha = 1$, and the vector correlation between points is expressed as $\phi\ (\alpha, \beta)$ where α is the horizontal separation and β is the vertical separation between points. The correlation $\phi\ (0, 0)$ was computed to be 10.40 and $\phi\ (-2, 2)$ was computed to be 5.67. These agree quite well with the values estimated for a larger ensemble of input noise where these values are easily seen to be 10 and 5, respectively. Other correlation values were found which were higher than the expected limiting values of 1, (square of the average value of the input), the highest being $\phi\ (2, 2) = 2.27$. These values ranged from .21 to 2.27 and were used in the simulation as is instead of assuming their limiting values. Thus the effect of having insufficient noise data to work from entered the simulation in a rather natural way.

Figure 12a is a picture of the noise input. It is normalized to the highest value, a 20 caused by two points overlapping. Figure 12b is the same picture with the addition of the target. Figure 12c is the signal plus noise picture after filtering using a 3 x 3 neighborhood in which the neighborhood values as found by the computer are

2	2.9	2.4
3.2	8.5	3.2
2.4	2.9	2

This seems intuitively correct. Values along the 135° where noise correlation is high are deemphasized.

Figure 12d is the same as 12c except that the neighborhood is 5 x 5 with values of

-1.5	-.6	-.5	-.4	-.5
0	.3	.5	.4	0
.5	1.0	2.8	1.0	.5
0	.4	.5	.3	0
-.5	-.4	-.5	-.6	-1.5

These neighborhood values again seem intuitively correct since they show deemphasis along the axis of the dumbbell noise.

The intensity over the target is seen to peak for both filters.

Figure 13 (Case 2) is the result of a second simulation of this filter using a different noise background. The noise in this case has amplitudes given by a rectangular probability density distribution of values between 0 and 9. The values at the points are completely uncorrelated with values at other points. This is a sample of the same noise used to generate the autocorrelation values with little fluctuation from the expected ratio $\phi(\alpha, \beta)/\phi(0, 0) = .7$, $\alpha \neq 0$, $\beta \neq 0$.

Figure 13a is the target plus noise normalized to the highest value. Figure 13b is the same picture filtered by a 3 x 3

neighborhood and 13c is the picture again but filtered by a 5 x 5 neighborhood.

The fact that the target loses its shape is somewhat deleterious to the humans' ability to find the target even when processed. This is because the human looks for structure perhaps without great regard to intensity.

3. REFERENCES

1. A. Feinstein, "Foundations of Information Theory", McGraw-Hill Book Co., New York, 1958, pp 26, 27.

2. Elias, Grey, and Robinson, Journal of the Opt. Soc. of America, Vol. 42, 1952, pp 127-134.

3. L. Brillouin, "Science and Information Theory", Academic Press, Inc., New York, 1956, p 97.

4. G. F. Aroyan, Proceedings of the I.R.E., September 1959, pp 1561-1568.

5. N. Wiener, "Extrapolation, Interpolation, and Smoothing of Stationary Time Series", Technology Press of M.I.T., Cambridge, and John Wiley & Sons, New York, 1949.

6. W. Meyer-Eppler, G. Darius, "Information Theory Third London Symposium", Butterworths Sci. Pub., London, 1956, p 34.

7. L. P. Horwitz, G. L. Shelton, Proceedings of the I.R.E., January 1961, pp 181, 182.

8. D. Middleton, "An Introduction to Statistical Communication Theory", McGraw-Hill Book Co., Inc., 1960, pp 773 - 810.

9. A. Wald, "Statistical Decision Functions", John Wiley & Sons, Inc., 1950.

APPENDIX A

Referring to equation (2) on page 19 the denominator can be written as

$$\overline{\left(\sum_i c_i n_i\right)^2} = \sum c_i c_j \varphi_{ij}$$

and the numerator as

$$\left[\sum_i c_i (s_i + \bar{n}_i)\right]^2 = \left(\sum_i c_i \sigma_i\right)^2$$

where $\varphi_{ij} = \overline{n_i n_j}$, $\sigma_i = s_i + \varphi$, and $\varphi = \bar{n}_i$.

Thus equations (3) on page 23 become

$$\left(\sum_{ij} c_i c_j \varphi_{ij}\right) 2\sigma_n \left(\sum_i c_i \sigma_i\right) - \left(\sum_i c_i \sigma_i\right)^2 2 \left(\sum_i c_i \varphi_{in}\right) = 0 \qquad (1')$$

for all values of n, or more simply

$$\sum_i c_i \varphi_{in} = a\sigma_n \qquad (2')$$

for all n, where a is independent of n. Notice that $\varphi_{ij} = \varphi_{ji}$ and so using standard matrix notation with $\underline{k}$ representing the vector solution of the c_i we obtain,

$$\Phi \underline{k} = a\, \underline{\sigma} \qquad (3')$$

corresponding to equation (4) on page 23.

Since equation (5) represents a critical point of R then under the assumptions listed there a necessary and sufficient condition for this to represent an absolute maximum is for

$R(\underline{k}+\underline{h}) - R(\underline{k}) = F(\underline{h})$ to be non positive for all vector displacements, $\underline{h}$. If $\underline{h}$ is parallel to $\underline{k}$ it is clear that $F(\underline{h}) = 0$ for the value of R is not sensitve to the undetermined scalar a in $\underline{k}$.

Substituting $\underline{k}$ into R we obtain

$$R(\underline{k}) = (a\,\underline{\sigma}'\,\underline{\Phi}^{-1}\underline{\sigma})^2 / a^2(\underline{\sigma}'\,\underline{\Phi}^{-1}\underline{\sigma}) = \underline{\sigma}'\,\underline{\Phi}^{-1}\underline{\sigma}. \quad (4')$$

The quantity $\underline{\sigma}'\,\underline{\Phi}^{-1}\underline{\sigma}$ will be denoted by α and since the positive definiteness of $\underline{\Phi}$ implies that of $\underline{\Phi}^{-1}$ we have $\alpha > 0$ for a non trivial target and background.

Next for $R(\underline{k}+\underline{h})$ we have

$$R(\underline{k}+\underline{h}) = (a\alpha+\underline{h}'\underline{\sigma})^2 / (a^2\alpha+2a\underline{h}'\underline{\sigma}+\underline{h}'\,\underline{\Phi}\,\underline{h}) \quad (5')$$

or letting $\underline{h}'\underline{\sigma} = \gamma(\underline{h})$ and $\underline{h}'\,\underline{\Phi}\,\underline{h} = \beta(\underline{h})$,

$$R(\underline{k}+\underline{h}) = (a\alpha+\gamma)^2 / (a^2\alpha+2a\gamma+\beta). \quad (6')$$

Thus

$$F(\underline{h}) = (\gamma^2-\alpha\beta) / (\alpha a^2+2\gamma a+\beta). \quad (7')$$

Considering the denominator of equation (7') as a function g(a) of a, its discriminant is four times the numerator. If $\gamma^2-\alpha\beta$ were positive for a single value of $\underline{h}$ then an a could be found such that $g(a) > 0$ so the $F(\underline{h}) > 0$ and $\underline{k}$ is not an absolute maximum. If $\alpha\beta-\gamma^2$ is non negative for all $\underline{h}$ then $g(a) > 0$ (except when $\alpha\beta-\gamma^2 = 0$ and then for only one value of a does $g(a) = 0$)

and $F(\underline{h})$ is non positive so that $\underline{k}$ is an absolute maximum.

Arranging terms

$$\alpha\beta - \gamma^2 = (\underline{\sigma}' \Phi^{-1} \underline{\sigma})(\underline{h}' \Phi \underline{h}) - (\underline{h}' \underline{\sigma})^2 = \underline{h}' \left(\alpha \Phi - \underline{\sigma}\, \underline{\sigma}'\right) \underline{h} \tag{8'}$$

so that the necessary and sufficient condition reduces to the positive semi-definiteness of the matrix $\alpha \Phi - \underline{\sigma}\, \underline{\sigma}'$ as stated on page 23.

APPENDIX B

In the variational approach to maximizing

$$R = \frac{\left(\int f(\underline{\xi})\,\sigma(\underline{\xi})\,d\underline{\xi}\right)^2}{\int f^2(\underline{\eta})\,\nu(\underline{\eta})\,d\underline{\eta}} \tag{9'}$$

we assume the existence of a maximizing function $\varphi(\underline{\xi})$ and then consider a family of "neighboring" functions $\varphi(\underline{\xi}) + \epsilon\,\delta(\underline{\xi})$, where ϵ is an arbitrary real value and $\delta(\underline{\xi})$ is an arbitrary function which vanishes on the boundary of the N-dimensional space. The boundary in the present case can be taken at infinity. If $\varphi(\underline{\xi}) + \epsilon\,\delta(\underline{\xi})$ is substituted for $f(\underline{\xi})$ in equation (9') the R becomes a function of ϵ and $\delta(\underline{\xi})$ and it can be shown that a necessary condition for a relative maximum of R is that $\partial R/\partial\epsilon = 0$ for all $\delta(\underline{\xi})$ and for $\epsilon = 0$. Thus we proceed with

$$R(\epsilon,\delta) = \frac{\left(\int\left[\varphi(\underline{\xi}) + \epsilon\,\delta(\underline{\xi})\right]\sigma(\underline{\xi})\,d\underline{\xi}\right)^2}{\int\left[\varphi(\underline{\eta}) + \epsilon\,\delta(\underline{\eta})\right]^2\nu(\underline{\eta})\,d\underline{\eta}} \tag{10'}$$

to obtain from $\left[\partial R/\partial\epsilon\right]_{\epsilon=0} = 0$ the following,

$$\left(\int\varphi^2\nu\right)\left(\int\varphi\sigma\right)\left(\int\delta\sigma\right) = \left(\int\varphi\sigma\right)\left(\int\varphi\delta\nu\right) \tag{11'}$$

Letting $\int\varphi\sigma = s$ and $\int\varphi^2\nu = t$ we have

$$\int s(t\sigma - s\varphi\nu)\,\delta = 0. \tag{12'}$$

The assumption that the integrand of equation (12') is continuous with δ arbitrary implies that

$$s(t\sigma - s\varphi\nu) = 0 \text{ for all } \underline{\xi}. \tag{13'}$$

Thus, with non trivial signal and noise, so that $s \neq 0$, $t \neq 0$ we have

$$\varphi(\underline{\xi}) = A\,\frac{\sigma(\underline{\xi})}{\nu(\underline{\xi})} \quad . \tag{14'}$$

General Dynamics/Astronautics,
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
(Contract AF 04(647)-644)
Unclassified report

Three spatial filters (linear, general statistical, and decision) are analyzed mathematically. The linear filter is simulated on the IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED

General Dynamics/Astronautics,
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
(Contract AF 04(647)-644)
Unclassified report

Three spatial filters (linear, general statistical, and decision) are analyzed mathematically. The IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED

General Dynamics/Astronautics
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
(Contract AF 04(647)-644)
Unclassified report

Three spatial filters (linear, general statistical, and decision are analyzed mathematically. The linear filter is simulated on the IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED

General Dynamics/Astronautics,
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
Contract AF 04(647)-644
Unclassified report

Three spatial filters (linear, general statistical, and decision) are analyzed mathematically. The IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED

General Dynamics/Astronautics,
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
(Contract AF 04(647)-644)
Unclassified report

Three spatial filters (linear, general statistical, and decision) are analyzed mathematically. The linear filter is simulated on the IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED

General Dynamics/Astronautics
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
(Contract AF 04(647)-644)
Unclassified report

Three spatial filters (linear, general statistical, and decision are analyzed mathematically. The linear filter is simulated on the IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED

General Dynamics/Astronautics,
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
(Contract AF 04(647)-644)
Unclassified report

Three spatial filters (linear, general statistical, and decision) are analyzed mathematically. The IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED

General Dynamics/Astronautics,
San Diego, California
SPACE-TIME SAMPLING AND FILTERING,
by P. W. Broome and W. D. Montgomery, May 1961. p. illus.
Exhibit "C"
Contract AF 04(647)-644
Unclassified report

Three spatial filters (linear, general statistical, and decision) are analyzed mathematically. The IBM 7090 with two-dimensional input and output.

UNCLASSIFIED

1. Polynomial filters
2. General statistical filter
3. Decision filter
4. Simulation

UNCLASSIFIED